



Bachelor's Thesis

to obtain the academic degree
— Bachelor of Science —
in Digital Business

at the Faculty of Business, Economics and Management Information Systems at the
University of Regensburg

Volatility Forecasts for Foreign Exchange Markets with Shrinkage Estimators

Submitted to Prof. Dr. Daniel Rösch
Chair of Statistics and Risk Management

Submitted by David Harrieder
Matr. Nr.: 2345350
Regensburgerstraße 8
84048 Mainburg
david.harrieder@stud.uni-regensburg.de

Submission date 05.08.2025

Abstract

Volatility forecasting in foreign-exchange (FX) markets is critical for risk management, derivative pricing, and portfolio allocation. This thesis investigates ...

Forecast accuracy is assessed via mean-squared error, mean-absolute error, QLIKE, and the Diebold–Mariano test for predictive superiority. ... The obtained results indicate that shrinkage methods, ...

Contents

List of Figures	iv
List of Tables	v
1 Introduction	1
2 Theoretical Foundations	1
2.1 Realized Variance	2
2.2 Forecasting Models	3
2.2.1 Heterogeneous Autoregressive (HAR) Model	3
2.2.2 Heterogeneous Exponential (HExp) Model	4
2.3 Shrinkage Estimators	5
2.3.1 Lasso-Regularized Extension Models	5
2.3.2 Cross-Validation for Lasso Model Selection	6
2.4 Forecasting Framework	7
2.4.1 Rolling-Window Framework	7
2.4.2 Forecast Evaluation Metrics	9
3 Foreign Exchange Markets Data	9
4 Empirical Analysis	12
4.1 Implementation	12
4.2 In-Sample Results	13
4.3 Out-of-Sample Results	14
5 Conclusion	18
6 Conclusion Brainstorm	19
6.1 Critics for Summary	19
Bibliography	20

List of Figures

1	Realized Variance and Kernel Density Estimate of RV	11
2	FX Spot Prices and Daily Returns	12
3	Lasso Regularization Paths	13
4	Relative QLIKE Performance of Lasso-Based Models	16
5	22-Day-Ahead Forecasts for USD/JPY	17
6	Coefficient Selection Patterns in the Lasso-HAR Model	17
7	Coefficient Selection Patterns in the Lasso-HExp Model	18

List of Tables

1	In-Sample Coefficient Estimates and Fit Metrics	14
2	Rolling-Window Out-of-Sample Forecast Evaluation	15
3	Diebold–Mariano Test Statistics	16

1 Introduction

Volatility forecasting lies at the heart of modern financial risk management, underpinning derivative pricing, dynamic hedging, and capital allocation decisions. In foreign-exchange (FX) markets, marked by deep liquidity, near-continuous trading, and swift reactions to macroeconomic and geopolitical developments, accurate estimates of return variance are indispensable. Market participants rely on such estimates not only to price currency options and manage risk exposures, but also to guide central-bank interventions and to craft algorithmic trading strategies that exploit transient volatility patterns. Episodes of market turbulence—ranging from the 2008 global financial crisis and the European sovereign-debt turmoil to the oil-price fluctuations of 2019–2020 and the COVID-19 pandemic—have repeatedly underscored the premium placed on robust measures of return variance. The academic literature has evolved in tandem with advances in data availability and computational tools. Early empirical research calibrated GARCH-type models (Bollerslev, 1986) to low-frequency returns, effectively capturing clustering and asymmetry in daily volatility dynamics. The proliferation of high-frequency transaction data later gave rise to realized variance (RV) estimators (Andersen and Bollerslev, 1998; Andersen, 2001; Andersen et al., 2003), which enabled more granular modeling of the latent return variance process. A particularly influential contribution is Corsi’s heterogeneous autoregressive (HAR) model (Corsi, 2008), which parsimoniously decomposes the persistence of realized variance across daily, weekly, and monthly horizons while retaining the appealing simplicity of ordinary least squares estimation.

Nevertheless, several limitations of the canonical HAR structure have spurred ongoing refinements. Its rigid horizon partitioning may fail to adapt when the underlying persistence shifts, and abrupt forecast revisions can occur when extreme observations enter or exit the moving averages (Audrino and Knaus, 2016; Audrino et al., 2017). Recent extensions address these concerns by integrating flexible lag selection and regularization techniques, such as the HAR(p) with data-driven shrinkage (Audrino et al., 2019), or by adopting smoother aggregation schemes like kernel-weighted and exponentially weighted HAR variants (Ghysels et al., 2019; Ding et al., 2021). Against this backdrop, the present study revisits the HAR framework for FX realized variance by systematically embedding shrinkage methods that allow the data to select relevant lags beyond the standard horizon sets. This approach aims to test whether relaxing the pre-specified structure can improve forecast accuracy, particularly in environments prone to regime shifts and structural breaks. In doing so, this work contributes to the growing body of research that combines the interpretability of HAR-type models with the flexibility of modern regularization techniques.

2 Theoretical Foundations

This section introduces the methodological foundations and modeling strategies applied throughout the analysis. As the basis for volatility modeling, realized variance is presented as a nonparametric estimator for the latent integrated variance, constructed from high-frequency financial data. Building on this foundation, two established forecasting models are outlined: the Heterogeneous Autoregressive (HAR) model and its exponential-weighted extension, the

Heterogeneous Exponential (HExp) model. To address potential issues of overfitting and high-dimensionality, both models are extended using Lasso-based shrinkage techniques, allowing for data-driven variable selection within a penalized estimation framework.

The subsequent forecasting procedure is based on a rolling-window framework designed to ensure strict out-of-sample prediction, followed by the introduction of forecast evaluation metrics commonly employed in volatility forecasting literature. Overall, the section is structured modularly: each modeling element is first introduced in isolation, before the components are combined into a coherent forecasting and evaluation framework. This structure aims to provide a clear conceptual understanding of the methodological approach before its empirical application.

2.1 Realized Variance

This subsection establishes the methodological and theoretical foundations for estimating latent integrated variance using high-frequency observations. It clarifies how continuous-time price dynamics give rise to observable proxies, which serve as the core inputs for the forecasting models developed in this study. Assume that the logarithmic price X_t of a financial asset evolves according to a continuous-time Itô semimartingale defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, given by

$$dX_t = \mu_t dt + \sigma_t dW_t, \quad (1)$$

where μ_t denotes a locally bounded, adapted drift process, σ_t is a right-continuous (càdlàg) stochastic volatility process, and W_t is a standard Brownian motion that is independent of σ_t . The main object of interest is the *integrated variance* over a single trading day, denoted IV_t , and defined as

$$\text{IV}_t = \int_{t-1}^t \sigma_s^2 ds, \quad (2)$$

which represents the total cumulative variance of the instantaneous returns dX_s within the interval from $t - 1$ to t (Barndorff-Nielsen and Shephard, 2002). In practice, the latent volatility path σ_t is not observable. However, the increasing availability of high-frequency financial data makes it possible to construct consistent nonparametric estimators of IV_t . The most widely used estimator is the *realized variance* (*RV*). Suppose that within a given trading day t , the log-price process X_t is observed at M equally spaced time points, which we denote $X_{t,0}, X_{t,1}, \dots, X_{t,M}$. The intraday returns are defined as $r_{t,i} = X_{t,i} - X_{t,i-1}$ for $i = 1, \dots, M$, where $r_{t,i}$ represents the log-return over the i -th interval of day t . Based on these high-frequency returns, the realized variance for day t is given by

$$\text{RV}_t = \sum_{i=1}^M r_{t,i}^2. \quad (3)$$

Under mild regularity conditions, as the sampling frequency increases (i.e., as $M \rightarrow \infty$ and the interval length $\Delta = 1/M \rightarrow 0$), the realized variance converges in probability to the integrated variance:

$$\text{plim}_{M \rightarrow \infty} \text{RV}_t = \text{IV}_t$$

(Andersen et al., 2003; Meddahi, 2002). In this paper, we use *realized variance* as our empirical proxy for the unobservable integrated volatility process. This approach is consistent with a

large body of empirical literature (e.g. Andersen et al., 2003; Barndorff-Nielsen, 2004; Corsi, 2008), as realized variance provides a simple, additive, and asymptotically unbiased estimate of the underlying daily volatility. It thus serves as a robust foundation for the forecasting models developed herein.

2.2 Forecasting Models

The following subsection outlines the two benchmark models underlying the forecasting framework: the Heterogeneous Autoregressive (HAR) model and its exponential-weighted extension, the Heterogeneous Exponential (HExp) model. Both specifications model volatility using structured combinations of lagged realized variance.

2.2.1 Heterogeneous Autoregressive (HAR) Model

The Heterogeneous Autoregressive (HAR) model, introduced by Corsi (2008), generalizes the Heterogeneous Autoregressive Conditional Heteroskedasticity (HARCH) framework of Müller et al. (1997) by representing realized variance (RV) as a linear combination of volatility components aggregated over daily, weekly, and monthly sampling intervals. This structure accounts for volatility clustering and long-range dependence, as evidenced in high-frequency financial time series (Andersen et al., 2003; Barndorff-Nielsen, 2004), while avoiding the complexity of fractional integration or other fully parametric long-memory specifications (Corsi, 2008). Formally, the realized variance at discrete time t , denoted RV_t , is specified as:

$$RV_t = \beta_0 + \beta_1 \overline{RV}_{t-1}^{(d)} + \beta_2 \overline{RV}_{t-1}^{(w)} + \beta_3 \overline{RV}_{t-1}^{(m)} + \varepsilon_t, \quad (4)$$

where $\beta_0 \in \mathbb{R}$ is the intercept parameter, $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)^\top \in \mathbb{R}^3$ denotes the vector of slope coefficients, and the error term $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ is assumed to be independently and identically distributed (i.i.d.) (Wooldridge, 2020). The lagged realized variance aggregates are defined as follows:

$$\begin{aligned} \overline{RV}_{t-1}^{(d)} &= RV_{t-1}, \\ \overline{RV}_{t-1}^{(w)} &= \frac{1}{5} \sum_{i=1}^5 RV_{t-i}, \\ \overline{RV}_{t-1}^{(m)} &= \frac{1}{22} \sum_{i=1}^{22} RV_{t-i}. \end{aligned} \quad (5)$$

In a general moving-average form:

$$\overline{RV}_{t-1}^{(i)} = \frac{1}{i} \sum_{j=1}^i RV_{t-j}, \quad (6)$$

where i denotes the aggregation horizon expressed in trading days ($i = 1$ for daily, $i = 5$ for weekly, $i = 22$ for monthly), and j is the summation index. Each aggregated component thus represents an equally weighted moving average of lagged realized variances (Corsi, 2008). The HAR specification captures heterogeneous trading horizons and temporal aggregation effects,

reflecting the behavior of both short-horizon and long-horizon market participants (Corsi, 2008). Crucially, the model retains analytical tractability: the parameters can be consistently and efficiently estimated via Ordinary Least Squares (OLS) under idealised assumptions (Wooldridge, 2020).

2.2.2 Heterogeneous Exponential (HExp) Model

The Heterogeneous Exponential (HExp) model, introduced by Bollerslev et al. (2018), generalizes the HAR model by replacing fixed-interval moving averages with normalized exponentially declining weights. Unlike the HAR specification, which aggregates realized variance (RV) over fixed, non-overlapping daily, weekly, and monthly windows, the HExp model utilizes Exponentially Weighted Moving Averages (EWMA), introducing smooth memory decay and attenuating the abrupt structural breaks inherent to fixed-window aggregations (Bollerslev et al., 2018). In its unrestricted form, the HExp model possesses theoretically infinite memory, which complicates empirical implementation and renders direct comparability with HAR-type models problematic (Bollerslev et al., 2018). Therefore, a truncated version is employed to restrict the memory span to a finite window, analogous to the maximum aggregation horizon in the HAR model, while maintaining exponential weighting.

For a decay parameter $\kappa \in (0, 1)$, the truncated exponentially weighted moving average over n trading days is specified as:

$$\text{TExpRV}_{t-1}^{(n)} = \sum_{i=1}^n \tilde{w}_i^{(\kappa)} \cdot RV_{t-i}, \quad \text{where } \tilde{w}_i^{(\kappa)} = \frac{(1 - \kappa)\kappa^{i-1}}{1 - \kappa^n}. \quad (7)$$

Here, $\tilde{w}_i^{(\kappa)}$ denotes normalized exponentially declining weights that assign higher importance to more recent observations, reflecting the empirical relevance of recency in volatility dynamics (Bollerslev et al., 2018). This weighting scheme captures the economic rationale of declining informational value for temporally distant observations while preserving long memory in a decaying functional form (Bollerslev et al., 2018). To ensure comparability with HAR-based specifications, the decay parameter κ is treated as a fixed constant throughout the analysis. This choice yields finite-memory, exponentially weighted averages consistent with the HAR aggregation horizons. Superscripts d , w , and m denote truncation over 1 (daily), 5 (weekly), and 22 (monthly) trading days, respectively:

$$\begin{aligned} \text{TExpRV}_{t-1}^{(d)} &= \text{TExpRV}_{t-1}^{(1)}, \\ \text{TExpRV}_{t-1}^{(w)} &= \text{TExpRV}_{t-1}^{(5)}, \\ \text{TExpRV}_{t-1}^{(m)} &= \text{TExpRV}_{t-1}^{(22)}. \end{aligned} \quad (8)$$

The resulting Truncated Heterogeneous Exponential model specifies the realized variance as:

$$RV_t = \beta_0 + \beta_1 \cdot \text{TExpRV}_{t-1}^{(d)} + \beta_2 \cdot \text{TExpRV}_{t-1}^{(w)} + \beta_3 \cdot \text{TExpRV}_{t-1}^{(m)} + \varepsilon_t, \quad (9)$$

where $\beta_0 \in \mathbb{R}$ is the intercept term, $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3)^\top \in \mathbb{R}^3$ is the vector of slope coefficients, and the error term $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ is assumed to be independently and identically distributed (i.i.d.) (Wooldridge, 2020). Parameter estimation proceeds OLS under classical linear regression

assumptions (Wooldridge, 2020). By integrating the concept of heterogeneous memory with the adaptive flexibility of exponential decay, while preserving aggregation horizon comparability, the truncated HExp model constitutes a structured yet flexible framework for modeling realized volatility (Bollerslev et al., 2018). For notational simplicity, the abbreviation HExp henceforth refers to the truncated version of the Heterogeneous Exponential model throughout this study.

2.3 Shrinkage Estimators

The following introduces shrinkage-based forecasting models, incorporating Lasso regularization, as well as the cross-validation procedure used for parameter selection and model estimation.

2.3.1 Lasso-Regularized Extension Models

To explore predictive structures beyond the fixed aggregation horizons used in classical HAR and HExp models—namely the 1-, 5-, and 22-day realized variance components—we extend these frameworks using the Least Absolute Shrinkage and Selection Operator (Lasso), introduced by Tibshirani (1996). This extension enables model flexibility by incorporating a richer set of lagged volatility predictors, including components located beyond the standard monthly window.

Lasso is an ℓ_1 -penalized regression method that combines coefficient shrinkage with variable selection, producing sparse representations focused on the most informative inputs (Hastie et al., 2009). This regularization approach is particularly well-suited to high-dimensional time series settings, where extensive lag structures may introduce overfitting, multicollinearity, or excess model variance.

Formally, we define a unified predictor notation $Z_{t-1}^{(i)}$, representing the lagged volatility inputs:

$$Z_{t-1}^{(i)} = \begin{cases} \bar{RV}_{t-1}^{(i)}, & \text{for Lasso-HAR,} \\ TExpRV_{t-1}^{(i)}, & \text{for Lasso-HExp,} \end{cases} \quad i = 1, \dots, p, \quad (10)$$

where p denotes the total number of candidate lagged volatility components. In the Lasso-HAR framework, $Z_{t-1}^{(i)}$ corresponds to discrete moving averages; in the Lasso-HExp framework, to truncated exponentially weighted averages with varying decay spans (Bollerslev et al., 2018).

The general linear forecasting model is specified as:

$$RV_t = \beta_0 + \sum_{i=1}^p \beta_{i,t}(\lambda) \tilde{Z}_{t-1}^{(i)} + \varepsilon_t, \quad \text{where } \varepsilon_t \sim \mathcal{N}(0, \sigma^2) \text{ i.i.d.} \quad (11)$$

(Wooldridge, 2020). Here, β_0 denotes the intercept term and $\beta(\lambda) = (\beta_1(\lambda), \dots, \beta_p(\lambda))^{\top} \in \mathbb{R}^p$ is the vector of slope coefficients estimated under a given regularization level. The predictors $\tilde{Z}_{t-1}^{(i)}$ are standardized versions of $Z_{t-1}^{(i)}$, calculated within the local estimation window \mathcal{R}_t ; details of this standardization process are formally presented in Equation (??) in Section 2.4.1.

Estimation proceeds by solving the ℓ_1 -penalized least squares optimization problem over the local estimation window \mathcal{R}_t :

$$\hat{\beta}_{i,t}(\lambda) = \arg \min_{\beta_0, \beta} \left\{ \sum_{t \in \mathcal{R}_t} \left(RV_t - \beta_0 - \sum_{i=1}^p \beta_i \tilde{Z}_{t-1}^{(i)} \right)^2 + \lambda \sum_{i=1}^p |\beta_i| \right\}, \quad \text{s.t. } \beta_i \geq 0 \forall i. \quad (12)$$

Here, $\lambda \geq 0$ is the regularization parameter controlling the degree of shrinkage (Tibshirani, 1996), and the intercept β_0 remains unpenalized to preserve identifiability. Additionally, a non-negativity constraint $\beta_i \geq 0$ is imposed to reflect the economic intuition that greater past volatility should not predict lower future risk (Eriksson et al., 2019).

The value of λ is selected adaptively at each forecast origin via time series cross-validation (see Section 2.3.2). Setting $\lambda = 0$ yields the ordinary least squares solution, while increasing λ induces sparser models, driving coefficients toward zero.

2.3.2 Cross-Validation for Lasso Model Selection

To select the regularization parameter λ in the high-dimensional Lasso forecasting model introduced in Section 2.3.1, we implement a time series cross-validation (TSCV) procedure tailored to the rolling-window framework described in Section 2.4.1. Within each estimation window $\mathcal{R}_t = \{RV_{t-w}, \dots, RV_{t-1}\}$ of length w , we apply an expanding-window cross-validation strategy that preserves temporal ordering and ensures strict out-of-sample prediction. This structure follows to a great extent the principles of blocked cross-validation for dependent data, particularly in its emphasis on preserving temporal ordering and avoiding leakage, while relying on the expanding-window logic implemented in `scikit-learn's TimeSeriesSplit` (Bergmeir and Benítez, 2012; Pedregosa et al., 2012).

The number of cross-validation folds is denoted by K , and the validation set size remains constant across folds, computed as $\lfloor w/(K+1) \rfloor$. For each fold $k \in \{1, \dots, K\}$, the training segment is defined as $\{RV_{t-w}, \dots, RV_{t_{\text{train}}^{(k)}}\}$ and the subsequent validation block as $\{RV_{t_{\text{train}}^{(k)}+1}, \dots, RV_{t_{\text{val}}^{(k)}}\}$, with $t_{\text{val}}^{(k)} \leq t-1$, ensuring that only data within \mathcal{R}_t are used for model selection.

For each candidate penalty parameter $\lambda \in \Lambda = \{\lambda_1, \dots, \lambda_L\}$, the Lasso model is trained on the standardized predictors $\tilde{Z}_{t-1}^{(i)}$ and targets RV_t observed within the training fold. The construction and standardization of $\tilde{Z}_{t-1}^{(i)}$ are defined in Section 2.4.1, based on localized inputs from \mathcal{R}_t .

The resulting coefficient estimates $\hat{\beta}_{i,t}^{(k)}(\lambda)$ are obtained via cyclic coordinate descent as detailed in Friedman et al. (2010), using warm starts, fixed-order soft-thresholding updates, and convergence criteria based on relative coefficient changes and duality gap thresholds.

Prediction accuracy is evaluated via mean squared error (MSE) over the validation block:

$$\text{MSE}_{\lambda}^{(k)} = \frac{1}{t_{\text{val}}^{(k)} - t_{\text{train}}^{(k)}} \sum_{t=t_{\text{train}}^{(k)}+1}^{t_{\text{val}}^{(k)}} \left(RV_t - \hat{RV}_t^{(k)}(\lambda) \right)^2, \quad (13)$$

where $\hat{RV}_t^{(k)}(\lambda)$ denotes the forecasted realized variance at time t , generated using the coefficients $\hat{\beta}_{i,t}^{(k)}(\lambda)$ estimated on the training set. These coefficients remain fixed throughout the validation

block, consistent with batch-style forecasting as implemented in `scikit-learn`'s (Pedregosa et al., 2012) `LassoCV`.

Validation errors are averaged across folds to compute the cross-validated objective:

$$\text{CV}(\lambda) = \frac{1}{K} \sum_{k=1}^K \text{MSE}_\lambda^{(k)}. \quad (14)$$

The optimal regularization parameter is then defined as

$$\lambda^* = \arg \min_{\lambda \in \Lambda} \text{CV}(\lambda), \quad (15)$$

and the model is retrained on the full estimation window \mathcal{R}_t using λ^* , yielding the final coefficients $\hat{\beta}_{i,t}(\lambda^*)$ for forecasting at origin t .

2.4 Forecasting Framework

This subsection outlines the forecasting procedure and evaluation methodology. First, the rolling-window forecasting framework used to generate strictly out-of-sample predictions is described. Subsequently, the forecast evaluation metrics employed to assess model performance are introduced.

2.4.1 Rolling-Window Framework

To generate strictly out-of-sample forecasts of cumulative realized variance, we employ a fixed-length rolling-window forecasting procedure (see, e.g., Bollerslev et al. (2016); Andersen et al. (2003)). Unlike recursive or expanding-window estimators, the rolling-window framework maintains adaptability by restricting the estimation sample to the most recent w observations, balancing bias-variance trade-offs under potential non-stationarity (?Tashman, 2000).

At each forecast origin t , the cumulative realized variance over a horizon of h trading days is defined as:

$$RV_t^{(h)} = \sum_{j=0}^{h-1} RV_{t+j}, \quad (16)$$

enabling direct modeling of multi-day volatility (Corsi, 2008).

Model parameters are re-estimated at each forecast origin using a localized estimation window of fixed size:

$$\mathcal{R}_t = \{RV_{t-w}, RV_{t-w+1}, \dots, RV_{t-1}\}, \quad (17)$$

ensuring strict chronological separation between training and forecasting, thus preventing look-ahead bias (Bergmeir et al., 2018).

Predictors follow the unified notation $Z_{t-1}^{(i)}$, introduced in Section 2.3.1, and are constructed exclusively from observations in \mathcal{R}_t . That is, for each $i = 1, \dots, p$, $Z_{t-1}^{(i)}$ implicitly depends on data up to time $t - 1$ within \mathcal{R}_t . This ensures temporal causality and compliance with rolling-window forecasting design (Tashman, 2000).

In HAR and HExp models (Sections 2.2.1 and 2.2.2), predictors are fixed to $i \in \{1, 5, 22\}$, corresponding to daily, weekly, and monthly aggregation spans (Corsi, 2008; Bollerslev et al.,

2018). In contrast, Lasso-based models include the full set $i = 1, \dots, p$, with $p \leq w$ to ensure identifiability and reduce estimation variance (?).

Before estimation, each predictor is standardized across \mathcal{R}_t individually:

$$\tilde{Z}_{t-1}^{(i)} = \frac{Z_{t-1}^{(i)} - \mu_{\mathcal{R}_t}^{(i)}}{\sigma_{\mathcal{R}_t}^{(i)}}, \quad (18)$$

where $\mu_{\mathcal{R}_t}^{(i)}$ and $\sigma_{\mathcal{R}_t}^{(i)}$ denote the sample mean and standard deviation of predictor i in \mathcal{R}_t . This standardization ensures scale-invariance of the penalty and numerical stability in optimization (Hastie et al., 2009).

Model estimation follows a general linear form:

$$RV_t^{(h)} = \beta_{0,t} + \sum_{i=1}^q \beta_{i,t}(\lambda) \tilde{Z}_{t-1}^{(i)} + \varepsilon_t, \quad (19)$$

where $q = 3$ for HAR and HExp models and $q = p$ for Lasso-based specifications. For HAR and HExp models, parameters are estimated via ordinary least squares (OLS) (Wooldridge, 2020), which corresponds to setting the regularization parameter to $\lambda = 0$. In contrast, Lasso-based models solve the ℓ_1 -penalized least squares problem from Equation (12), where the optimal penalty level λ^* is re-tuned within each estimation window via localized time series cross-validation (Section 2.3.2).

The forecast of cumulative realized variance over horizon h at time t is then:

$$\widehat{RV}_t^{(h)} = \hat{\beta}_{0,t} + \sum_{i=1}^q \hat{\beta}_{i,t}(\lambda^*) \cdot \tilde{Z}_{t-1}^{(i)}, \quad (20)$$

where λ^* is the cross-validated regularization level and $\hat{\beta}_{i,t}(\lambda^*)$ are the corresponding coefficient estimates obtained from \mathcal{R}_t .

Forecast origins progress in horizon-aligned steps of h days:

$$\mathcal{F}_h = \{t_0, t_0 + h, t_0 + 2h, \dots\}, \quad \text{with } t_0 = w + 1, \quad (21)$$

ensuring non-overlapping forecast targets and avoiding serial correlation in forecast errors (Tashman, 2000; ?).

Collecting all forecasts yields the sequence:

$$\left\{ \widehat{RV}_t^{(h)} : t \in \mathcal{F}_h \right\}, \quad (22)$$

where each prediction is computed solely from information contained in its respective \mathcal{R}_t .

Finally, for Lasso-based models, the active predictor set at each forecast origin is defined as:

$$\hat{S}_t(\lambda^*) = \left\{ i \in \{1, \dots, p\} : \hat{\beta}_{i,t}(\lambda^*) \neq 0 \right\}, \quad (23)$$

facilitating post-estimation analysis of variable relevance and dynamic feature selection (?).

2.4.2 Forecast Evaluation Metrics

Forecast evaluation is essential for assessing the empirical performance of volatility forecasting models (Patton, 2011). In this study, forecast accuracy is measured by comparing model-generated forecasts to subsequently realized volatility, using three complementary evaluation metrics: the quasi-likelihood loss (QLIKE), the mean squared error (MSE), and the root mean squared error (RMSE). These metrics capture distinct facets of predictive performance and are widely adopted in volatility forecasting research (Patton and Sheppard, 2015; ?; ?; ?).

Let \mathcal{F}_h denote the set of forecast origins for horizon h , as defined in Section 2.4.1, and $|\mathcal{F}_h|$ the total number of forecasts generated. At each forecast origin $t \in \mathcal{F}_h$, the model generates a forecast $\widehat{RV}_t^{(h)}$ for cumulative realized variance, which is compared to the realized outcome $RV_t^{(h)}$. All metrics are computed as averages over the entire out-of-sample forecast sequence for each horizon.

The quasi-likelihood (QLIKE) loss serves as the primary evaluation criterion (Patton, 2011). Designed to accommodate the positive skewness and measurement noise typical of realized variance, the QLIKE loss is defined as:

$$\text{QLIKE}_h = \frac{1}{|\mathcal{F}_h|} \sum_{t \in \mathcal{F}_h} \left(\frac{RV_t^{(h)}}{\widehat{RV}_t^{(h)}} - \log \frac{RV_t^{(h)}}{\widehat{RV}_t^{(h)}} - 1 \right). \quad (24)$$

This likelihood-based loss penalizes underprediction more severely than overprediction, reflecting the asymmetric cost structure inherent in volatility forecasting (Patton, 2011; ?). QLIKE is considered robust in the presence of noise-contaminated volatility proxies, making it particularly suitable for evaluating realized variance forecasts in financial applications (Patton, 2011).

Complementing the likelihood-based evaluation, forecast accuracy is also assessed using the mean squared error (MSE), defined as:

$$\text{MSE}_h = \frac{1}{|\mathcal{F}_h|} \sum_{t \in \mathcal{F}_h} \left(RV_t^{(h)} - \widehat{RV}_t^{(h)} \right)^2, \quad (25)$$

which measures the average squared forecast deviation. To improve interpretability, the root mean squared error (RMSE) is computed as:

$$\text{RMSE}_h = \sqrt{\text{MSE}_h}, \quad (26)$$

restoring the forecast error to the original scale of realized variance. Both MSE and RMSE are sensitive to large forecast errors due to their quadratic loss structure and remain standard benchmarks in time-series forecast evaluation (??).

3 Foreign Exchange Markets Data

This section describes the high-frequency foreign exchange (FX) market data used in the empirical analysis. The dataset contains tick-level and minute-by-minute records for three major currency pairs: EUR/USD, GBP/USD, and USD/JPY. These pairs are among the most actively traded instruments globally and provide a solid basis for analyzing intraday volatility. The

data was obtained from HistData.com (<https://www.histdata.com>) and covers intraday price movements at one-minute intervals. Trading occurs continuously from Sunday 5:00 p.m. to Friday 4:59 p.m. Eastern Standard Time (EST). No observations are recorded during the weekend closure from Friday evening to Sunday evening. Each trading day begins at 5:00 p.m. EST of the previous calendar day and ends at 4:59 p.m. EST of the stated date. For instance, data labeled 2012-01-03 includes observations from 2012-01-02 17:00:00 to 2012-01-03 16:59:00 EST. This rolling-day convention ensures that volatility estimates reflect actual trading patterns rather than arbitrary calendar cutoffs. To avoid holiday effects, all observations for New Year's Day (January 1st) and Christmas Day (December 25th) have been excluded. The final sample spans from February 3, 2012 to December 30, 2022, capturing a decade of trading conditions that include calm periods as well as episodes of increased uncertainty. In total, 2853 data points are being used in the further analysis.

Figure ?? displays the daily realized variance (RV) of the EUR/USD, GBP/USD, and USD/JPY exchange rates over the period from 2012 to 2022, complemented by kernel density estimates (KDEs). Despite broad similarities, each series exhibits distinct distributional characteristics and volatility regimes. The EUR/USD market shows a relatively balanced volatility profile over the sample, with moderate dispersion around low base levels. However, pronounced spikes emerge between 2015 and 2017, peaking in mid-2016—likely reflecting the market turmoil surrounding the Brexit referendum and European banking concerns. Additional episodes of elevated variance appear briefly in early 2020 at the onset of the COVID-19 pandemic and again from 2022 onward, potentially linked to monetary policy uncertainty and geopolitical instability. The corresponding KDE reveals a smooth, moderately right-skewed tail, with a density concentration near zero and rare but sizable excursions up to 0.0008. GBP/USD displays structurally lower baseline volatility but features sharp episodic surges. The most prominent spike aligns with June 2016, the month of the Brexit vote, followed by additional increases in late 2016 and early 2020. From mid-2022, volatility remains moderately elevated. The KDE for GBP/USD is highly leptokurtic, with an extremely narrow modal region and a long upper tail extending to 0.005, indicative of rare yet extreme fluctuations. USD/JPY mirrors some of the temporal dynamics of EUR/USD but with greater heterogeneity. Elevated volatility is observed in 2013 amid aggressive monetary easing under Japan's Abenomics policies, and again between 2015 and 2017, peaking around early 2016. Short-lived spikes in early 2019 and early 2020 correspond to flash crashes and pandemic-induced shocks, respectively. Volatility also rises from mid-2022, consistent with global interest rate realignments. The KDE profile is right-skewed but smoother than GBP/USD, with an upper support reaching 0.0025. Overall, all three currencies exhibit periods of regime shifts driven by macroeconomic or geopolitical shocks. While EUR/USD and USD/JPY exhibit comparable volatility structure and persistence, GBP/USD stands out due to its relatively stable baseline and rare but dramatic variance spikes.

Figure 2 provides an overview of the spot price trajectories and corresponding daily returns for EUR/USD, GBP/USD, and USD/JPY over the 2012–2022 period. EUR/USD shows a moderate downward trend, with notable return spikes around mid-2016 (Brexit) and early 2020 (COVID-19). GBP/USD exhibits a more pronounced depreciation, accompanied by heightened volatility during and after the Brexit vote. USD/JPY displays a structural appreciation until

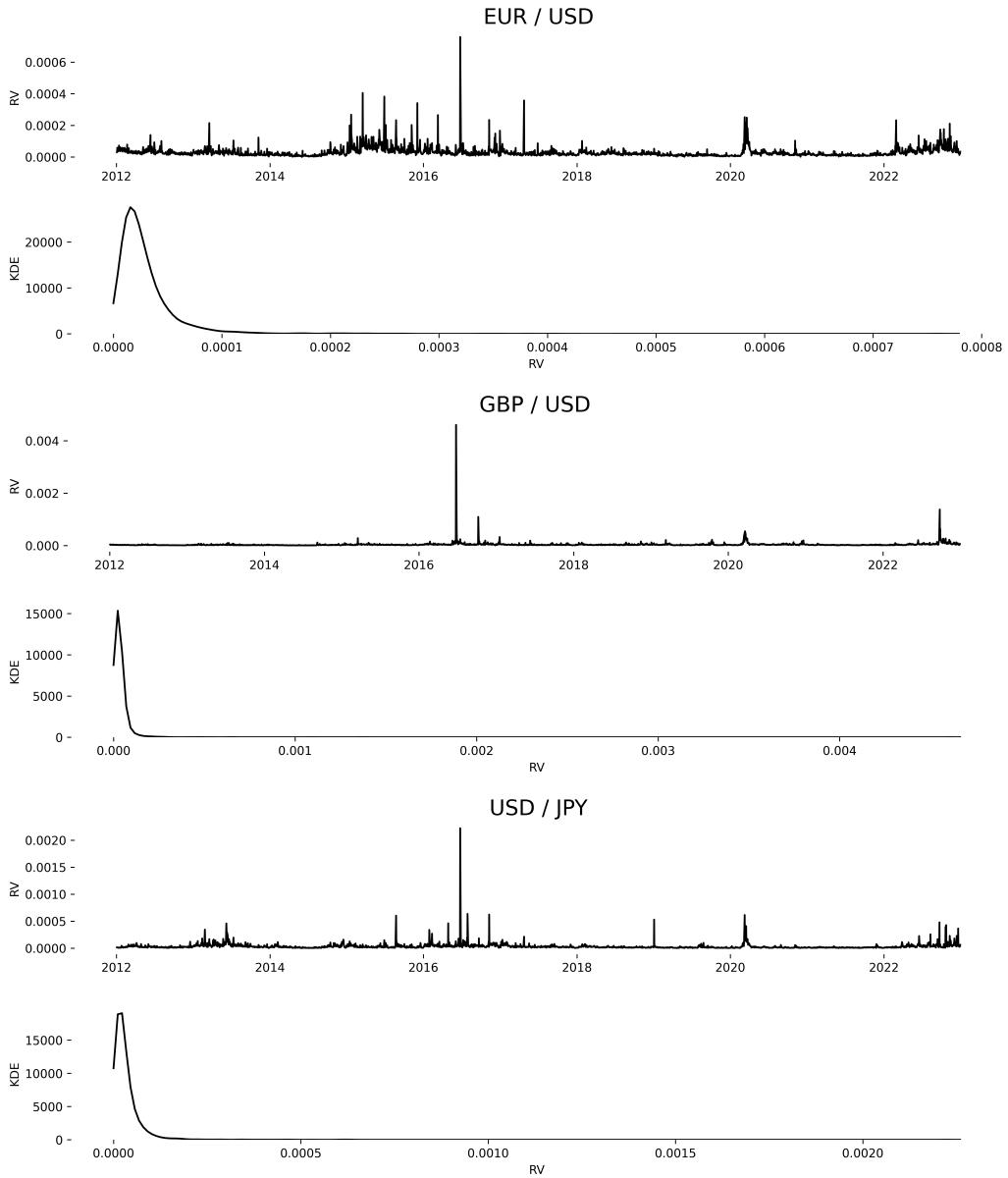


Figure 1: Realized Variance and Kernel Density Estimate of RV. This figure displays the time series of daily realized variance (RV_t) and the corresponding kernel density estimates for the three major FX pairs: EUR/USD (top), GBP/USD (middle), and USD/JPY (bottom), over the sample period from January 2012 to December 2022.

2015, stabilizes, and then accelerates again post-2021. Return volatility is most pronounced for GBP/USD, while EUR/USD and USD/JPY show sharper but more isolated return shocks. These plots serve to contextualize the volatility characteristics of the analyzed currency pairs and offer a coherent visual foundation for the subsequent modeling efforts.

Since we want the purest view of the model, we intentionally do not perform any sort of outlier removal.

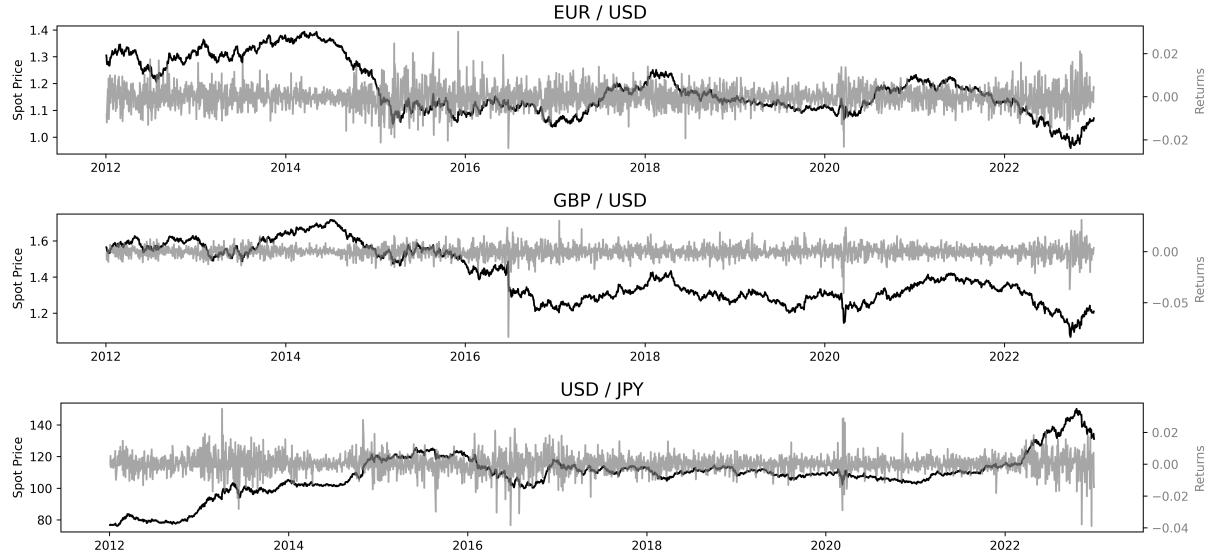


Figure 2: FX Spot Prices and Daily Returns. Daily spot prices (black, left axis) and log-returns (grey, right axis) for EUR/USD, GBP/USD, and USD/JPY over the period 2012–2022.

4 Empirical Analysis

Despite the empirical appeal of HAR-type volatility models, their rigid horizon partitioning can limit responsiveness to structural changes in volatility dynamics (Audrino and Knaus, 2016; Audrino et al., 2017). To address this, we incorporate Lasso-regularized extensions of the HAR and HExp models, enabling flexible lag selection over an extended predictor set. This approach builds on findings by Ding et al. (2021), who show that lagged volatility beyond the conventional 22-day window retains predictive power. Consequently, we implement four models for realized variance RV_t : HAR, HExp, Lasso-HAR, and Lasso-HExp, regarding the three major currency pairs, described in Section 3. In the following, we go through the practical implementation, then present a short in-sample evaluation, followed by an extensive out-of-sample evaluation, in which we also examine the selected lag structure in more detail.

4.1 Implementation

The forecasting setup follows the rolling-window framework introduced in Section 2.4.1, with forecast horizons $h \in \{1, 5, 22\}$ and a fixed estimation window size of $w = 1000$ trading days.

The benchmark HAR and HExp models utilize fixed predictors with lags $i \in \{1, 5, 22\}$, while Lasso-HAR and Lasso-HExp models extend the predictor set to $p = 100$ daily lags. For the HExp and Lasso-HExp models, the exponential decay parameter is set to $\kappa = 0.94$, following the RiskMetrics Framework by J.P. Morgan (*RiskMetrics™ — Technical Document*, 1996).

In both Lasso extensions, the regularization parameter λ is selected dynamically at each forecast origin using localized time series cross-validation, as detailed in Section 2.3.2. This is operationalized via the `LassoCV` routine from `scikit-learn` (Pedregosa et al., 2012), which plays a central role in the forecasting pipeline. For each forecast origin $t \in \mathcal{F}_h$, `LassoCV` fits the model on standardized predictors $\tilde{Z}_{t-1}^{(i)}$ using 10 expanding-window folds ($K = 10$) within the estimation window \mathcal{R}_t . The regularization grid Λ consists of 100 logarithmically spaced values

between λ_{\max} and $\lambda_{\min} = \lambda_{\max} \times 10^{-3}$, recomputed in each window to maintain adaptiveness to the evolving data distribution. Here, λ_{\max} is the smallest value that shrinks all coefficients to zero. Convergence is monitored through duality gap thresholds, with a fixed tolerance level of 10^{-3} and a maximum of 100,000 iterations. For consistency and comparability, the benchmark models are also estimated using the same standardized predictor set $\tilde{Z}_{t-1}^{(i)}$, although restricted to the HAR-specific subset $i \in \{1, 5, 22\}$. Forecast performance is assessed using the accuracy metrics introduced in Section 2.4.2.

Figure 3 illustrates the effect of the lasso procedure: it plots the regularization paths of all lag coefficients as λ varies over a logarithmic scale. Redundant predictors are progressively eliminated as penalization intensifies, while informative lags remain active.

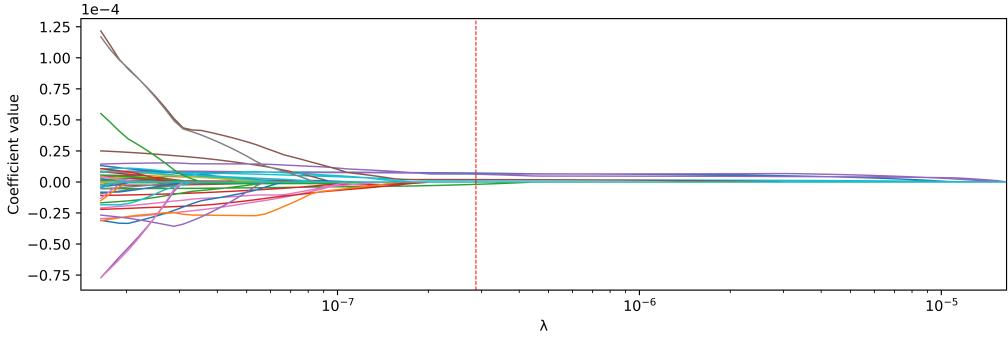


Figure 3: Lasso Regularization Paths. Coefficient trajectories for the Lasso-HAR model with forecast horizon $h = 1$ and maximum lag order $p = 100$. Each curve represents the evolution of a lag coefficient as the penalty parameter λ varies on a logarithmic scale. The vertical dashed line indicates the optimal regularization parameter $\lambda = 2.87418e-07$, selected via time-series cross-validation. Retained coefficients correspond to the HAR moving-average form (6) with lags $i = 1, 5, 10, 45, 76$. Results are based on an exemplary out-of-sample forecast for January 20, 2017.

All models are implemented in **Python**, utilizing **pandas** (McKinney, 2010), **numpy** (Harris et al., 2020), and **statsmodels** (Seabold and Perktold, 2010), alongside **scikit-learn** (Pedregosa et al., 2012).

Because the flexible Lasso-based models incorporate lagged realized variances up to 100 days, the rolling training window for all models is initialized on May 22, 2012, ensuring sufficient historical depth for parameter estimation. For consistency, this start date is also applied to the baseline HAR and HExp models. Consequently, the first out-of-sample forecast is issued on May 1, 2016, marking the beginning of the out-of-sample forecast evaluation period.

4.2 In-Sample Results

Although the main goal of this study is to evaluate out-of-sample forecast performance, it is also informative to examine the in-sample results. Table 1 summarizes the in-sample fit metrics and coefficient estimates (by lag) for each model and each currency pair. Following Bollerslev et al. (2016) and related empirical work (e.g., Audrino and Knaus, 2016; Ding et al., 2021), we report in-sample fits as a diagnostic benchmark to illustrate the magnitude, sign, and statistical significance of the estimated lag coefficients. For each currency pair, the in-sample estimation uses the entire available sample with a one-day-ahead forecasting scheme. Because the Lasso-based

models include up to 100 lags, the first valid forecast date is December 22, 2012. All parameters are estimated under the same data preprocessing as in the rolling-window exercise.

Table 1: In-Sample Coefficient Estimates and Fit Metrics. Reported are coefficient estimates and model fit metrics (MSE, RMSE, R^2 , and adjusted R^2) for the HAR, HExp, Lasso-HAR (L-HAR), and Lasso-HExp (L-HExp) specifications, applied to the estimation of daily realized variance for the EUR/USD, GBP/USD, and USD/JPY exchange rates. Asterisks (*) denote statistical significance at the 5% level; boldface indicates the best-performing model in each panel. Gray shading highlights benchmark HAR lags (1, 5, and 22 days) for reference.

Lag / Metric	EUR/USD				GBP/USD				USD/JPY			
	HAR	HExp	L-HAR	L-HExp	HAR	HExp	L-HAR	L-HExp	HAR	HExp	L-HAR	L-HExp
MSE	7.5290e-10	7.5350e-10	7.4712e-10	7.5005e-10	9.3004e-09	9.2913e-09	9.2542e-09	9.2738e-09	3.2693e-09	3.2637e-09	3.2595e-09	3.2699e-09
RMSE	2.7439e-05	2.7450e-05	2.7333e-05	2.7387e-05	9.6439e-05	9.6391e-05	9.6199e-05	9.6300e-05	5.7178e-05	5.7129e-05	5.7092e-05	5.7183e-05
R^2	0.2987	0.2982	0.3041	0.3014	0.0861	0.0870	0.0907	0.0888	0.1319	0.1334	0.1345	0.1317
adj. R^2	0.2980	0.2974	0.3026	0.3002	0.0851	0.0860	0.0887	0.0865	0.1309	0.1324	0.1373	0.1353
Intercept	2.9022e-05*	2.9022e-05*	2.9022e-05*	2.9022e-05*	3.7067e-05*	3.7067e-05*	3.7067e-05*	3.7067e-05*	3.3121e-05*	3.3121e-05*	3.3121e-05*	3.3121e-05*
1	5.2640e-06*	4.8574e-06*	5.2783e-06*	4.8811e-06*	1.4094e-05*	1.3222e-05*	1.3079e-05*	1.2386e-05*	8.1184e-06*	7.5160e-06*	7.1989e-06*	6.9679e-06*
2	—	—	3.7753e-06*	2.0299e-06	—	—	—	—	—	—	1.4744e-06	9.7765e-07
4	—	—	—	—	—	—	—	1.0148e-05	8.1909e-06*	—	—	8.2247e-07
5	9.2060e-06*	7.6931e-06*	7.1425e-06*	6.6469e-06*	1.2061e-05*	8.8429e-06*	2.0402e-08	—	8.2520e-06*	5.0491e-06*	—	—
6	—	—	1.7909e-06	—	—	—	—	—	—	—	4.9624e-06	4.3406e-06
10	—	—	2.0831e-06	—	—	—	—	—	—	—	2.8828e-06	1.8838e-06
14	—	—	—	—	—	—	9.6415e-06*	8.1714e-06	—	—	—	—
22	5.4834e-06*	6.9906e-06*	—	—	9.9023e-06*	1.2891e-05*	—	—	1.0057e-05*	1.3023e-05*	—	—
25	—	—	—	—	—	—	—	—	—	—	2.5919e-06	—
41	—	—	—	—	—	—	—	—	—	—	2.7767e-06	-2.2335e-06
45	—	—	2.0831e-06	—	—	—	—	—	—	—	—	—
75	—	—	—	—	—	—	4.8430e-06	-7.6831e-06	—	—	—	—
76	—	—	2.0831e-06	—	—	—	—	—	—	—	—	—
98	—	—	5.2434e-07	-6.1176e-04	—	—	—	2.4650e-04	—	—	3.0921e-06	1.2800e-05
99	—	—	—	—	—	—	—	-2.0303e-03	—	—	—	—
100	—	—	—	6.1797e-04	—	—	7.2403e-07	1.7981e-03	—	—	—	—

In summary, the flexible Lasso-HAR model (L-HAR) delivers the best in-sample fit across all three currencies, achieving the highest R^2 . The one-day lag coefficient consistently emerges as the strongest predictor of next-day variance, whereas longer lags receive progressively smaller weights or are eliminated by the Lasso, underscoring the dominant role of recent volatility. USD/JPY exhibits a lower overall R^2 than EUR/USD and GBP/USD, reflecting its more idiosyncratic volatility dynamics.

4.3 Out-of-Sample Results

We next evaluate the rolling-window out-of-sample forecasting performance of all models across the three currency pairs and different forecast horizons.

Table 2 reports the out-of-sample forecast performance for all models in terms of QLIKE, MSE, and RMSE. Following the volatility forecast comparison literature (Patton, 2011), we treat QLIKE as the primary metric due to its robustness to measurement error in the volatility proxy and its greater penalty for under-prediction, using realized variance as the proxy.

At the 1-day horizon, Lasso regularization is particularly effective. The Lasso-HExp model achieves the lowest QLIKE for both EUR/USD and GBP/USD, indicating improved probabilistic calibration that is valuable in risk-sensitive applications. Moreover, Lasso-HExp remains competitive in terms of MSE and RMSE, suggesting that it improves distributional fit while preserving point forecast accuracy. By the 5-day horizon, the advantage of Lasso diminishes. EUR/USD still exhibits a marginal QLIKE improvement under Lasso-HExp, but any gains in MSE or RMSE disappear and can even reverse. For USD/JPY, the base HExp model outperforms its

Table 2: Rolling-Window Out-of-Sample Forecast Evaluation. Forecast performance is evaluated for the HAR, HExp, Lasso-HAR, and Lasso-HExp models across the EUR/USD, GBP/USD, and USD/JPY exchange rates, using rolling-window estimation. Forecasts target the daily realized variance $\widehat{RV}_t^{(h)}$ at horizons $h = 1$, 5, and 22 trading days. Performance is assessed using QLIKE, MSE, and RMSE as loss metrics. For each currency-horizon combination, the lowest (i.e., best) value per metric is shown in **bold**.

Model	EUR/USD			GBP/USD			USD/JPY		
	QLIKE	MSE	RMSE	QLIKE	MSE	RMSE	QLIKE	MSE	RMSE
1-day									
HAR	0.1391	7.0991e-10	2.6644e-05	0.1590	4.9326e-08	2.2209e-04	0.2499	4.7723e-09	6.9082e-05
HExp	0.1394	7.0962e-10	2.6639e-05	0.1578	4.9978e-08	2.2356e-04	0.2472	4.7537e-09	6.8947e-05
Lasso HAR	0.1370	7.0771e-10	2.6603e-05	0.1578	4.6797e-08	2.1633e-04	0.2519	4.7154e-09	6.8669e-05
Lasso HExp	0.1362	7.0438e-10	2.6540e-05	0.1570	4.8217e-08	2.1958e-04	0.2532	4.7150e-09	6.8666e-05
5-day									
HAR	0.0477	3.5901e-09	5.9918e-05	0.0597	3.5540e-08	1.8852e-04	0.1175	1.6368e-08	1.2794e-04
HExp	0.0480	3.5631e-09	5.9692e-05	0.0580	3.2827e-08	1.8118e-04	0.1146	1.5964e-08	1.2635e-04
Lasso HAR	0.0456	3.8584e-09	6.2116e-05	0.0651	3.8818e-08	1.9702e-04	0.1239	1.7440e-08	1.3206e-04
Lasso HExp	0.0447	3.6162e-09	6.0135e-05	0.0575	3.4338e-08	1.8531e-04	0.1245	1.6590e-08	1.2880e-04
22-day									
HAR	0.0280	3.2534e-08	1.8037e-04	0.0438	4.0609e-07	6.3725e-04	0.0768	1.2041e-07	3.4700e-04
HExp	0.0284	3.1977e-08	1.7882e-04	0.0402	2.6893e-07	5.1859e-04	0.0734	1.0860e-07	3.2955e-04
Lasso HAR	0.0264	3.6863e-08	1.9200e-04	0.0488	3.8405e-07	6.1972e-04	0.0860	1.5377e-07	3.9213e-04
Lasso HExp	0.0254	3.2272e-08	1.7964e-04	0.0408	2.9264e-07	5.4096e-04	0.0836	1.3143e-07	3.6254e-04

regularized counterpart on all metrics at $h = 5$. This pattern is even more pronounced at the 22-day horizon: The unpenalized HExp yields superior distributional fit, as indicated by lower QLIKE values, and greater point forecast accuracy (lower MSE/RMSE) than the Lasso-based models, particularly for the GBP/USD and USD/JPY series. While the Lasso-HAR model does achieve marginal QLIKE improvements in isolated cases (e.g., EUR/USD at 5 days), it generally fails to outperform its baseline HAR counterpart. In most settings, it yields higher MSE and RMSE, suggesting that the sparsity constraint offers limited value when applied to an already parsimonious lag structure. This is particularly evident at longer horizons, where predictive signal appears distributed across more extended lag ranges.

Figure 4 summarizes these findings visually. The plot shows that only EUR/USD consistently benefits from Lasso regularization, with both Lasso-HAR and Lasso-HExp yielding relative QLIKE ratios below 1 across all horizons. By contrast, for GBP/USD and USD/JPY this ratio exceeds 1 at longer horizons, implying reduced alignment between forecasted and realized variance for the Lasso-based forecasts. This divergence highlights that a uniform penalization strategy can be detrimental in regimes with more persistent volatility.

Residuals across all models and forecast horizons exhibit clear deviations from normality, as indicated by high skewness and excess kurtosis. These fat-tailed distributions are particularly evident in short-horizon forecasts, where large outliers in realized volatility likely contribute to the breakdown of Gaussian assumptions. A comprehensive residual analysis lies outside the scope of this paper. Table 3 reports Diebold–Mariano (Diebold and Mariano, 1995) test statistics comparing each baseline model (HAR, HExp) to its Lasso-regularized counterpart (Lasso-HAR, Lasso-HExp) across currencies and forecast horizons. The test is based on squared forecast error differentials (MSPE differences), with HAC standard errors estimated via the Newey–West

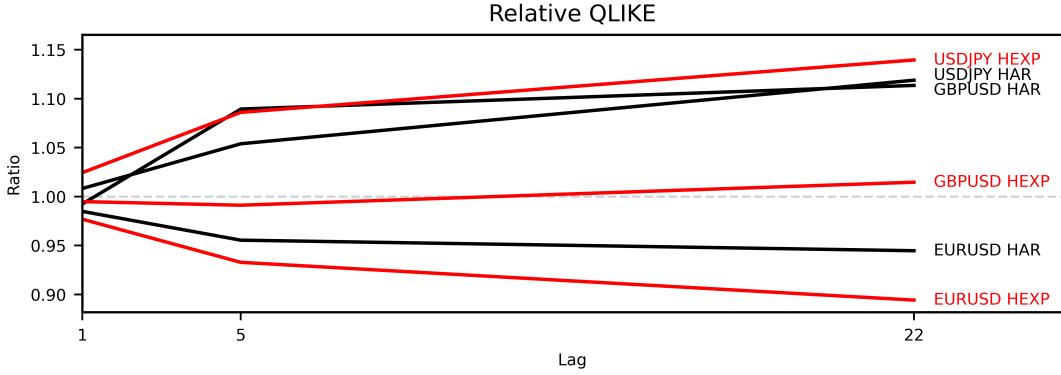


Figure 4: Relative QLIKE Performance of Lasso-Based Models. Each ratio is $Q\text{LIKE}_{\text{Lasso}} / Q\text{LIKE}_{\text{Base}}$, normalized such that 1 corresponds to equal performance. Values below 1 indicate superior performance by the Lasso model compared to the equivalent base model. HExp Models are marked in red. The dashed line marks parity (ratio = 1).

method using a lag length equal to the respective forecast horizon (Newey and West, 1987). Positive values indicate that the Lasso model outperforms its baseline counterpart in terms of forecast accuracy.

Table 3: Diebold–Mariano Test Statistics. Test statistics from the Diebold–Mariano test using HAC standard errors (Newey–West estimator), comparing baseline models to their Lasso-regularized counterparts across forecast horizons $h = 1, 5, 22$. Positive values indicate improved predictive accuracy from Lasso regularization. p -values are reported in parentheses; * and ** denote significance at the 5% and 1% levels, respectively.

Horizon	Comparison	EUR/USD	GBP/USD	USD/JPY
1-day	HAR vs. Lasso-HAR	0.5504 (0.5821)	1.0585 (0.2900)	1.0618 (0.2885)
	HExp vs. Lasso-HExp	2.1401* (0.0325)	1.0682 (0.2856)	1.0673 (0.2860)
5-day	HAR vs. Lasso-HAR	-1.3773 (0.1693)	-1.3885 (0.1659)	-1.5333 (0.1261)
	HExp vs. Lasso-HExp	-0.4597 (0.6460)	-0.9511 (0.3422)	-1.4198 (0.1566)
22-day	HAR vs. Lasso-HAR	-1.7568 (0.0829)	0.8524 (0.3966)	-2.0064* (0.0483)
	HExp vs. Lasso-HExp	-0.2487 (0.8042)	-0.9301 (0.3552)	-2.8335** (0.0059)

The DM tests reveal no consistent winner between the baseline and Lasso-augmented models. The DM statistics vary in sign and magnitude across horizons and currencies, and most are not significantly different from zero. There is no systematic evidence favoring either the traditional or the regularized approach in terms of overall forecast accuracy. Figure 5 provides the statistically significant difference (Diebold–Mariano test, 1% level) in 22-day-ahead USD/JPY volatility forecasts, highlighting the Lasso model’s tendency to underpredict during volatile periods like 2016–2017.

To gain deeper insight into the Lasso models’ behavior, Figures 6 and 7 show kernel density scatter plots of the selected lag orders $\hat{S}_t(\lambda^*)$ versus their coefficient magnitudes (vertical axis) for the Lasso-HAR and Lasso-HExp models, respectively. Each figure displays results for $h = 1, 5$, and 22 days, allowing us to examine which lag lengths are chosen and how their influence varies with the forecast horizon.

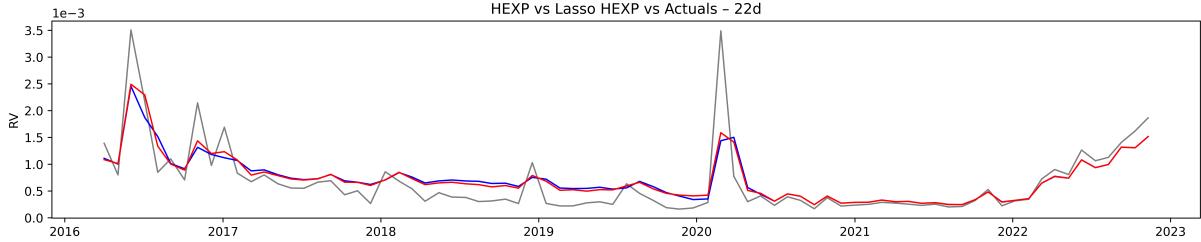


Figure 5: 22-Day-Ahead Forecasts for USD/JPY. Rolling-window forecasts from the truncated heterogeneous exponential (HExp) model (red) and its Lasso-regularized (Lasso-HExp) variant (blue), with realized volatility shown in gray. Forecast horizon: 22 trading days.

For the Lasso-HAR model, the selection pattern is dominated by short-term lags across all horizons (Figure 6).

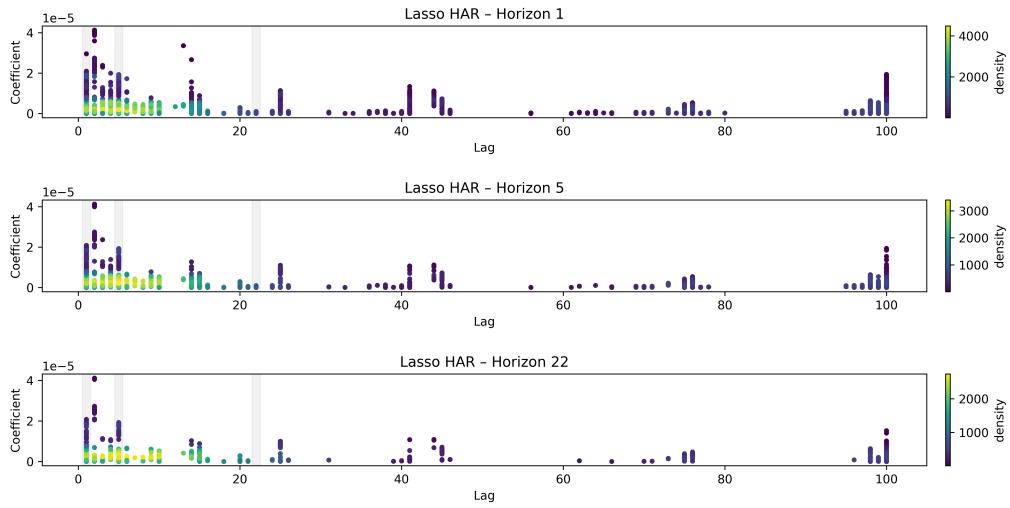


Figure 6: Coefficient Selection Patterns in the Lasso-HAR Model. Kernel density scatter plots of selected lag orders versus estimated coefficient magnitudes across forecast horizons $h = 1, 5$, and 22 . Each point corresponds to a retained lag term in the Lasso-HAR specification. Gray shading highlights benchmark lags (1, 5, and 22 days) from the classical HAR model for reference.

In all three panels, the Lasso-HAR consistently picks the most recent lags (within the past 10 days), with the largest coefficients typically on $\ell = 1, 2$, and 5 . Beyond the 10-day mark, only a few additional lags receive nonzero weights (notably around $\ell \approx 13, 25, 45$, and 75), and their coefficients are relatively small. Interestingly, the coefficient magnitudes tend to rise again at the very longest lags: large values often occur at $\ell = 100$, suggesting that the model occasionally assigns weight to the furthest lag in the pool.

The Lasso-HExp model shows a slightly different selection pattern (Figure 7).

Like Lasso-HAR, the Lasso-HExp puts most emphasis on the shortest lags (1, 2, and 5 days). However, beyond those, the selected lags form clear clusters rather than isolated points. In particular, there is a pronounced cluster of nonzero coefficients around $\ell = 25$, another dense region in the $\ell \approx 40\text{--}50$ range, and a broader selection of lags around $\ell = 75$. The coefficient magnitude in Lasso-HExp increases more gradually toward the end of the lag spectrum, near $\ell = 98$.

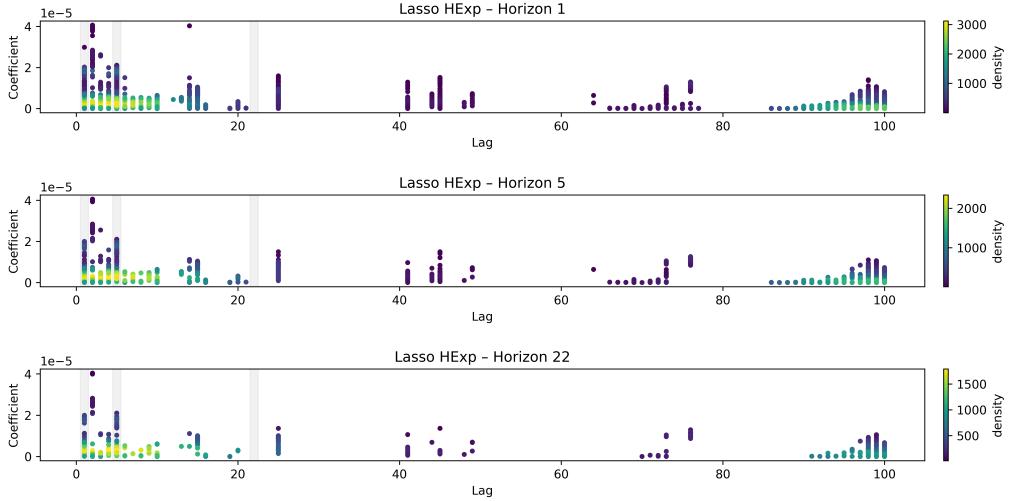


Figure 7: Coefficient Selection Patterns in the Lasso-HExp Model. Kernel density scatter plots of selected EWMA spans versus estimated coefficient magnitudes across forecast horizons $h = 1, 5$, and 22 . Each point represents a non-zero coefficient selected by the Lasso regularization. Gray shading marks benchmark HAR lags (1, 5, and 22 days) for comparison.

5 Conclusion

Across all models and horizons, these results illustrate the nuanced effects of Lasso regularization on volatility forecasting. Shrinkage tends to improve forecast performance with respect to distributional accuracy (lower QLIKE) at the shortest horizon, particularly for EUR/USD, where volatility is highly noisy. However, these benefits diminish as the horizon increases. For more persistent markets and longer forecast horizons, the unpenalized models frequently yield better point forecast accuracy in terms of MSE and RMSE. Overall, the efficacy of the Lasso is highly regime-dependent.

Finally, the selected lag structures reveal that both Lasso-HAR and Lasso-HExp concentrate on very recent volatility but also include some longer-term lags. This suggests that predictive information in realized volatility spans multiple time scales. The Lasso-HExp model, owing to its exponential smoothing design, selects groups of contiguous lags (forming clusters with smoothly varying coefficients), whereas the Lasso-HAR model picks a more disparate set of lags. These differences notwithstanding, neither approach categorically outperforms the other in out-of-sample terms, underscoring that while flexible regularization can yield gains in certain conditions, it must be applied with attention to the underlying volatility regime.

6 Conclusion Brainstorm

6.1 Critics for Summary

- Employ other Estimations than standard OLS, e.g. RR, WLS
- using k-fold cross validation (e.g. bergmeier)
- using not only one step ahead cross validation but according to horizon
- Other Transformations, e.g. Box-Cox, log-transformation,
- Value at Risk Evaluation und Interpretation
- Mehrere Variablen noch verwenden, bspw. RVN, RVP, dret, etc.
- Freijeiro-Gonzalez et al., A critical review of Lasso and its derivatives for variable selection under dependence among covariates
- Filterung von Ausreißern,
- Vollständige QLIKE Formel ?
- Residuen verringern durch Cochrane-Orcutt Transformation oder ARIMA, GARCH oder HAR-GARCH

Zu guter Letzt wird man seine Arbeit noch kurz und prägnant zusammenfassen wollen. Was sind die zentralen Erkenntnisse und „take aways“?

Eventuell kann man ganz zum Schluss auch noch weitere anschließende Forschungsfragen für die Zukunft skizzieren.

Bibliography

- Andersen, Torben G (2001), ‘The distribution of realized stock return volatility’, *Journal of Financial Economics* **61**(1), 43–76. Publisher: Elsevier BV.
URL: <https://linkinghub.elsevier.com/retrieve/pii/S0304405X01000551>
- Andersen, Torben G. and Tim Bollerslev (1998), ‘Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts’, *International Economic Review* **39**(4), 885.
URL: <https://www.jstor.org/stable/2527343?origin=crossref>
- Andersen, Torben G., Tim Bollerslev, Francis X. Diebold and Paul Labys (2003), ‘Modeling and Forecasting Realized Volatility’, *Econometrica* **71**(2), 579–625.
URL: <http://doi.wiley.com/10.1111/1468-0262.00418>
- Audrino, Francesco, Chen Huang and Ostap Okhrin (2019), ‘Flexible HAR model for realized volatility’, *Studies in Nonlinear Dynamics & Econometrics* **23**(3), 20170080.
URL: <https://www.degruyter.com/document/doi/10.1515/snnde-2017-0080/html>
- Audrino, Francesco, Lorenzo Camponovo, Constantin Roth and Department of Economics, University of St. Gallen, Bodanstrasse 6, 9000 St.Gallen, Switzerland (2017), ‘Testing the Lag Structure of Assets’ Realized Volatility Dynamics’, *Quantitative Finance and Economics* **1**(4), 363–387.
URL: <http://www.aimspress.com/article/10.3934/QFE.2017.4.363>
- Audrino, Francesco and Simon D. Knaus (2016), ‘Lassoing the HAR Model: A Model Selection Perspective on Realized Volatility Dynamics’, *Econometric Reviews* **35**(8-10), 1485–1521.
URL: <http://www.tandfonline.com/doi/full/10.1080/07474938.2015.1092801>
- Barndorff-Nielsen, O. E. (2004), ‘Power and Bipower Variation with Stochastic Volatility and Jumps’, *Journal of Financial Econometrics* **2**(1), 1–37.
URL: <https://academic.oup.com/jfec/article-lookup/doi/10.1093/jjfinec/nbh001>
- Barndorff-Nielsen, Ole E. and Neil Shephard (2002), ‘Econometric Analysis of Realized Volatility and its Use in Estimating Stochastic Volatility Models’, *Journal of the Royal Statistical Society Series B: Statistical Methodology* **64**(2), 253–280.
URL: <https://academic.oup.com/jrssb/article/64/2/253/7098420>
- Bergmeir, Christoph and José M. Benítez (2012), ‘On the use of cross-validation for time series predictor evaluation’, *Information Sciences* **191**, 192–213.
URL: <https://linkinghub.elsevier.com/retrieve/pii/S0020025511006773>
- Bergmeir, Christoph, Rob J. Hyndman and Bonsoo Koo (2018), ‘A note on the validity of cross-validation for evaluating autoregressive time series prediction’, *Computational Statistics & Data Analysis* **120**, 70–83.
URL: <https://linkinghub.elsevier.com/retrieve/pii/S0167947317302384>
- Bollerslev, Tim (1986), ‘Generalized autoregressive conditional heteroskedasticity’, *Journal of Econometrics* **31**(3), 307–327.
URL: <https://linkinghub.elsevier.com/retrieve/pii/0304407686900631>

Bollerslev, Tim, Andrew J. Patton and Rogier Quaedvlieg (2016), ‘Exploiting the errors: A simple approach for improved volatility forecasting’, *Journal of Econometrics* **192**(1), 1–18.
URL: <https://linkinghub.elsevier.com/retrieve/pii/S0304407615002584>

Bollerslev, Tim, Benjamin Hood, John Huss and Lasse Heje Pedersen (2018), ‘Risk Everywhere: Modeling and Managing Volatility’, *The Review of Financial Studies* **31**(7), 2729–2773.
URL: <https://academic.oup.com/rfs/article/31/7/2729/5001472>

Corsi, F. (2008), ‘A Simple Approximate Long-Memory Model of Realized Volatility’, *Journal of Financial Econometrics* **7**(2), 174–196.
URL: <https://academic.oup.com/jfec/article-lookup/doi/10.1093/jjfinec/nbp001>

Diebold, Francis X. and Roberto S. Mariano (1995), ‘Comparing Predictive Accuracy’, *Journal of Business & Economic Statistics* **13**(3), 253–263.
URL: <http://www.tandfonline.com/doi/abs/10.1080/07350015.1995.10524599>

Ding, Yi, Dimos Kambouroudis and David G. McMillan (2021), ‘Forecasting realised volatility: Does the LASSO approach outperform HAR?’, *Journal of International Financial Markets, Institutions and Money* **74**, 101386.
URL: <https://linkinghub.elsevier.com/retrieve/pii/S1042443121001050>

Eriksson, Anders, Daniel P. A. Preve and Jun Yu (2019), ‘Forecasting Realized Volatility Using a Nonnegative Semiparametric Model’, *Journal of Risk and Financial Management* **12**(3), 139. Publisher: MDPI AG.
URL: <https://www.mdpi.com/1911-8074/12/3/139>

Friedman, Jerome, Trevor Hastie and Robert Tibshirani (2010), ‘Regularization Paths for Generalized Linear Models via Coordinate Descent’, *Journal of Statistical Software* **33**(1).
URL: <http://www.jstatsoft.org/v33/i01/>

Ghysels, Eric, Alberto Plazzi, Rossen Valkanov, Antonio Rubia Serrano and Asad Dossani (2019), ‘Direct Versus Iterated Multi-Period Volatility Forecasts: Why MIDAS Is King’, *SSRN Electronic Journal*. Publisher: Elsevier BV.
URL: <https://www.ssrn.com/abstract=3326606>

Harris, Charles R., K. Jarrod Millman, Stéfan J. Van Der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. Van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández Del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke and Travis E. Oliphant (2020), ‘Array programming with NumPy’, *Nature* **585**(7825), 357–362.
URL: <https://www.nature.com/articles/s41586-020-2649-2>

Hastie, Trevor, Robert Tibshirani and Jerome Friedman (2009), *The Elements of Statistical Learning*, Springer Series in Statistics, Springer New York, New York, NY.
URL: <http://link.springer.com/10.1007/978-0-387-84858-7>

McKinney, Wes (2010), Data Structures for Statistical Computing in Python, Austin, Texas, pp. 56–61.

URL: <https://doi.curvenote.com/10.25080/Majora-92bf1922-00a>

Meddahi, Nour (2002), ‘A theoretical comparison between integrated and realized volatility’, *Journal of Applied Econometrics* **17**(5), 479–508.

URL: <https://onlinelibrary.wiley.com/doi/10.1002/jae.689>

Müller, Ulrich A., Michel M. Dacorogna, Rakhal D. Davé, Richard B. Olsen, Olivier V. Pictet and Jacob E. Von Weizsäcker (1997), ‘Volatilities of different time resolutions — Analyzing the dynamics of market components’, *Journal of Empirical Finance* **4**(2-3), 213–239.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S0927539897000078>

Newey, Whitney K. and Kenneth D. West (1987), ‘A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix’, *Econometrica* **55**(3), 703. Publisher: JSTOR.

URL: <https://www.jstor.org/stable/1913610?origin=crossref>

Patton, Andrew J. (2011), ‘Volatility forecast comparison using imperfect volatility proxies’, *Journal of Econometrics* **160**(1), 246–256. Publisher: Elsevier BV.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S030440761000076X>

Patton, Andrew J. and Kevin Sheppard (2015), ‘Good Volatility, Bad Volatility: Signed Jumps and The Persistence of Volatility’, *Review of Economics and Statistics* **97**(3), 683–697.

URL: <https://direct.mit.edu/rest/article/97/3/683-697/58249>

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Andreas Müller, Joel Nothman, Gilles Louppe, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot and Édouard Duchesnay (2012), ‘Scikit-learn: Machine Learning in Python’. Publisher: arXiv Version Number: 4.

URL: <https://arxiv.org/abs/1201.0490>

RiskMetrics™ — Technical Document (1996), Technical Document, J.P. Morgan, New York, NY.

URL: <http://www.jpmorgan.com/RiskManagement/RiskMetrics/RiskMetrics.html>

Seabold, Skipper and Josef Perktold (2010), statsmodels: Econometric and statistical modeling with python.

Tashman, Leonard J. (2000), ‘Out-of-sample tests of forecasting accuracy: an analysis and review’, *International Journal of Forecasting* **16**(4), 437–450.

URL: <https://linkinghub.elsevier.com/retrieve/pii/S0169207000000650>

Tibshirani, Robert (1996), ‘Regression Shrinkage and Selection Via the Lasso’, *Journal of the Royal Statistical Society Series B: Statistical Methodology* **58**(1), 267–288.

URL: <https://academic.oup.com/jrsssb/article/58/1/267/7027929>

Wooldridge, Jeffrey M. (2020), *Introductory econometrics: a modern approach*, seventh edition edn, Cengage Learning, Boston, MA.

Erklärung

Ich habe die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Die Arbeit wurde bisher an keiner anderen Hochschule zur Erlangung eines akademischen Grades eingereicht. Die vorgelegten Druckexemplare und die dem Prüfer/der Prüferin zur Verfügung gestellte elektronische Version (PDF-Datei) der Arbeit sind identisch. Von den in §27 Abs. 6 BPO 2021 vorgesehenen Rechtsfolgen habe ich Kenntnis.

Regensburg, 22.07.2025